



# gbif-indexingtoolkit

The GBIF Harvesting and Indexing Toolkit (HIT)

Search projects

[Project Home](#)

[Downloads](#)

**Wiki**

[Issues](#)

[Source](#)

Export to GitHub

Search

Current pages



for

Search

## GeneralSpecs

General Specifications

Updated Sep 17, 2010 by [kyle.br...@gmail.com](mailto:kyle.br...@gmail.com)

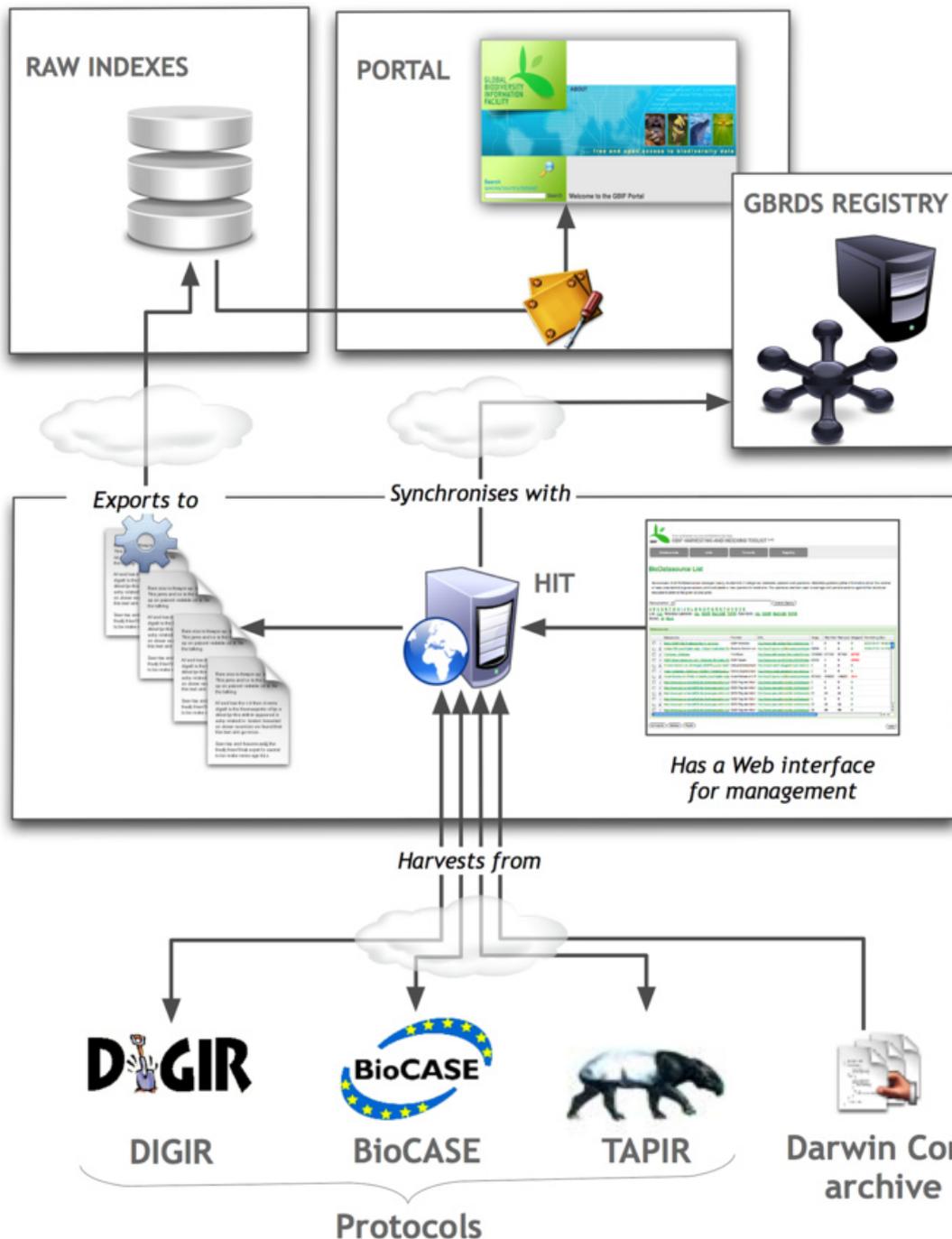
### General Specifications

The GBIF Harvesting and Indexing Toolkit (HIT) is an open source Java-based web application. The following document explains answers to these questions:

- [What protocols is the HIT capable of supporting?](#)
- [What data encodings are supported for each protocol?](#)
- [What are the outputs from harvesting and how can it be customised?](#)
- [What are the outputs from indexing and how can it be customised?](#)

The following diagram is a nice overview of the HIT's specifications:

# Overview of HIT Specifications



## What protocols is the HIT capable of supporting?

Support for the three most commonly used protocols (and their most popular versions and flavours) for publishing occurrence data has been added, meaning that currently it can harvest data being exposed via:

- [DiGIR 1.0](#) & [MaNIS DiGIR 1.0](#)

► [Sign in to add a comment](#)

- [TAPIR 1.0](#)

Additional support has also been added for the new [Darwin Core Archive standard](#), making data available for upload via a single dump. Built in such a way that it is fully extensible, the HIT allows support for additional protocols to be easily incorporated at a later time. For information on how to do so, please see the section entitled [Adding a New Protocol with Predefined Templates and Customisations](#) document.

## What data encodings are supported for each protocol?

Where applicable, a given protocol can allow the end-user to define how their XML requests and responses are encoded. Therefore, within a